Cheatsheet

HL Crowell[1,2], P-L Germain[1,2], and MD Robinson[1,2]

[1]Institute for Molecular Life Sciences, University of Zurich
[2]SIB Swiss Institute of Bioinformatics, Zurich

[BC]2 workshop | September 9th, 2019
"Analysis of multi-sample multi-group scRNA-seq data"

QC & filtering with `scater`

`addQCPerCell/Feature`	Compute cell-/gene/level QC metrics & add them to a SCE
`isOutlier`	Determine outliers based on median absolute deviation (MAD)

Integration & clustering with `Seurat`

`CreateSeuratObject`	Contruct a `SeuratObject` (SO)
`SplitObject`	Split a SO by some factor (e.g., sample ID)
`NormalizeData`	Normalization of count data
`FindVariableFeatures`	Identify HVGs based on the mean-variance relationship
`ScaleData`	Scale & center individual features
`FindIntegrationAnchors`	Identification of integration anchors
`IntegrateData`	Dataset integration using a pre-compute anchorset
`RunX`	Dimension reduction (`X = PCA, TSNE, UMAP, ...`)
`FindNeighbors`	Shared Nearest Neighbor (SNN) graph construction
`FindClusters`	Cluster identification based on a pre-computed SNN graph

DS Analysis with `muscat`

`prepSCE`	SCE reformatting for `muscat`
`aggregateData`	Aggregation of single-cell to pseudobulk data (e.g., sum of counts by cluster-sample)
`pbDS`	Sample-level DS analysis with `edgeR, limma, DESeq2`
`mmDS`	Cell-level DS analysis using mixed models
`calcExprFreqs`	Compute expression frequencies by sample and/or group
`resDS`	Reformat DS results into wide- or tidy-table format

Visualization with `scater` & `muscat`

`plotX`	Reduced dimension plot (`X = PCA, TSNE, UMAP, ...`)
`plotExpression`	Violin plots against categorical covariate (e.g., sample ID)
`pbMDS`	Pseudobulk-level multidimensional scaling plot
`pbHeatmap`	Heatmap of cluster-sample pseudobulks
`upset`	Upset plot, e.g., intersecting DS genes across clusters (`UpSetR` package)